# SHAPE ANALYSIS USING FOURIER DESCRIPTORS

*A Thesis Submitted*
in Partial Fulfilment of the Requirements
for the Degree of

MASTER OF TECHNOLOGY

*by*

V. DAMODAR BHAT

*to the*

DEPARTMENT OF ELECTRICAL ENGINEERING
INDIAN INSTITUTE OF TECHNOLOGY KANPUR
MARCH, 1986

CERTIFICATE

Certified that the work entitled 'SHAPE ANALYSIS USING FOURIER DESCRIPTORS' by V. Damodar Bhat has been carried out under our supervision and has not been submitted elsewhere for a degree.

( H.C. Karnick )
Lecturer
Department of Computer Science
Indian Institute of Technology
KANPUR

( P.R.K. Rao )
Professor
Department of Elect. Engg.
Indian Institute of Technology
KANPUR

ACKNOWLEDGEMENTS

It is with great pleasure that I record my thanks to my thesis supervisors, Dr. P.R.K. Rao and Dr. H.C. Karnick, not only for those absorbing and diverse discussions we had on Monday afternoons, but also for being accessible at all times of the day and night.

To my friends Anup Gogoi and Rakesh Saxena, my debt is too great to be repaid ever. Thanks are also due all the happy-go-lucky C-toppers and my other friends in Hall IV.

Finally, I must congratulate Mr. Rawat for doing an excellent job of typing the manuscript.

IIT, Kanpur
March, 1986

V. Damodar Bhat

TABLE OF CONTENTS


| Chapter | | Page |
|---|---|---|
| 1 | Introduction | 1 |
| | 1.1 Statement of the problem | 2 |
| | 1.2 Definition of shape | 4 |
| | 1.3 Some criteria for representation of shape | 6 |
| | 1.4 Algorithms for shape analysis | 7 |
| | 1.5 Plan of the thesis | 8 |
| 2 | Shape Analysis and Fourier Representation | 9 |
| | 2.1 Introduction | 9 |
| | 2.2 Fourier Descriptors: A Review | 9 |
| | 2.3 The Method | 15 |
| | 2.3.1 Fourier Descriptors of a plane closed contour | 15 |
| | 2.3.2 Elliptic properties of the Fourier coefficients | 18 |
| | 2.3.3 Rotation of the contour | 21 |
| | 2.3.4 Uniform scaling | 23 |
| | 2.3.5 Normalization of the Fourier descriptors | 24 |
| | 2.3.6 Some remarks | 26 |
| 3 | Classification of Shapes using Bending-Energy and Matched-Filtering | 27 |
| | 3.1 Introduction | 27 |
| | 3.2 Classification of shape using Bending-energy | 28 |
| | 3.3 Classification of shape using Matched-Filter | 36 |

| Chapter | | Page |
|---|---|---|
| 4 | Experiments, Results and Discussion | 40 |
| | 4.1 Introduction | 40 |
| | 4.2 The Global versus Local Nature of the Fourier representation | 40 |
| | 4.3 Classification based on Bending-Energy | 44 |
| | 4.4 Classification backed on Matched - Filtering | 60 |
| 5 | Conclusion and Further Work | 62 |
| | References | 66 |
| | Appendix | |

# LIST OF FIGURES


| Fig. No. | Caption | Page |
|---|---|---|
| 3.2.1 | An Elastic Bar under Deformation | 29 |
| 3.2.2 | Calculation of Bending Energy of a shape | 31 |
| 3.2.3 | A shape and its mirror image | 37 |
| 4.2.1 | Illustrating the Role of Lower and Higher Harmonics | 42 |
| 4.2.2 | Illustrating the Role of Lower and Higher Harmonics | 43 |
| 4.3.1 | Three Shapes and their Bending Energies | 45 |
| 4.3.2 | Bending Energies of three variants of B | 47 |
| 4.3.3 | 51 Variants of 'C' arranged in Descending Order of Bending Energy | 48 |
| 4.3.4 | A Shape and its Mirror Image | 59 |

## ABSTRACT

The broad aim of this thesis is to inquire into the nature of the Fourier series representation for shape. Specially, the emphasis is upon arriving at a suitable basis for classifying shapes which are closely related; the ease with which we can accomplish this task is an indication of the ability of the representation to capture fine details of the shape and to dissociate them from the less varying or gross features of the shape. We used two methods to classify shapes. In the first, the concept of bending energy stored in an elastic rod which is deformed to the shape in question is used. The expression for bending energy corresponds to the integral for the curvature over the entire boundary and is therefore intuitively satisfying. The second method treats the descriptors as parameters in a filter and shapes are classified by matching the parameters of the target with the known parameters as is done for matched filters in signal processing.

# CHAPTER 1

## INTRODUCTION

One of the main areas of interest in artificial intelligence is the use of representations for problem solving. A representation for an object is a symbolic abstraction of the object in the real world. In applied mathematics, for example, numbers and operations on numbers are used to represent the physical world and the actions in it. The main reason why one chooses to deal with representations rather than with the actual objects these representations stand for is that they bring important information to the foreground, and hence make it easy for us to solve the problem in the representation world. The nature of the problem dictates the choice of a particular representation; in general, the same object can have different representations emphasising different aspects of the object in the real world. For example, the so-called intrinsic representation (curvature vs. arc-length) for the boundary of an object brings out the curvedness or wiggliness of the boundary. The medial axis transform representation, on the other hand, brings out the correspondence of the opposite sides of a boundary, thus emphasizing width perception relating boundary arcs all along parallel sides.

The existence of many representations for an object immediately poses many important questions such as the extent to which a representation corresponds with the real world, of how powerful the representation is and what class of objects are representable.

It was this desire to probe into the nature of the representation that led to the present work. To be more specific, a pattern recognition system with Fourier descriptors being used as a representation for a 2D-shape had been built by Karnick [4 ]. The present thesis was conceived of as being complementary to the above work in as much as we decided to investigate the adequacy of the Fourier series representation for describing shapes-specifically those of hand written characters - in more detail.

## 1.1 STATEMENT OF THE PROBLEM:

Conventional pattern recognition is concerned more with classification among objects which are generally quite different from each other than with the particulars of differentiation between objects whose general shape is similar. In the recognition of hand written characters, for example, the problem of reliably distinguishing between two characters, say, an A and B has, understandably perhaps,

received much more attention than that of being able to distinguish between the natural variants of a character. The emphasis has generally been on interclass classification rather than intraclass separation.

The main concern of this thesis will be with the problem of distinguishing between 2D-shapes (we will clarify soon what is meant by shape) which, though they resemble each other in a global (overall) sense are neverthless, recognized by us as being different from each other because there are significant local variations between them. Our contention is that any scheme for representing shape must be able to perform, over and above interclass classification, finer classification within a class itself. In other words, we are concerned with the sensitivity of the representation to finer details. The problem before us then, is to attempt to achieve as fine a distinction between related shapes as possible within the framework of whatever representation we might choose to use for analyzing shape. Thus, in another and related sense, we shall inquire into the ability of a representation to capture the gross and fine features of a shape, and the degree to which they are decoupled from each other.

We shall, in this thesis, restrict ourselves to the domain of hand-written characters - as people would naturally write them. There are two excellent reasons for such a choice. For one, there exists a huge data base of hand-written characters collected by Karnick[4] and hence the time-consuming and tedious task of obtaining input data can be forgone. Moreover, we can match the results obtained against the judgement of a human observer, which in the final analysis, is what counts.

## 1.2 DEFINITION OF SHAPE:

It might be well to clarify our ideas as to what we mean by shape, and put it in mathematical terms. Since we are concerned only with 2D-shapes, we shall restrict ourselves to the Euclidean plane, $R^2$. The Euclidean plane $R^2$ is a two-dimensional vector space over the real numbers. Transformations of $R^2$ which preserve all ratios of distances are called similarities. In $R^2$, a similarity is either a translation, rotation about some origin, multiplication by a constant of all distances from the origin (uniform scaling), or one of these followed by a reflection.

We shall start with the notion of a boundary or outline. An outline or boundary will be defined as a smooth closed curve in $R^2$ i.e. a map of a part of the real line into $R^2$

which is twice differentiable 'almost everywhere' except at a finite number of points. Any outline in $R^2$ is mapped by the similarity transformations onto a four parameter family of other outlines. We can translate it to any origin (which is two degrees of freedom), rotate it (one degree of freedom) or alter its scale (one d.f.). Two outlines are equivalent if one can be mapped onto the other by a similarity transformation in $R^2$. The similarity transformations form a group - the so-called equiform group - and hence we may construct equivalence classes of outlines which are collections of all the outlines that can be mapped onto one another by similarity transformations. A shape in $R^2$ can now be defined as an equivalence class, under the equiform group, of outlines in $R^2$. Intuitively, we can consider a shape to be an outline which does not contain any information about position, scale and orientation in the coordinate frame. A shape representation or measurement is defined as a function on some domain into the real line which is the <u>same</u> for all elements of an equivalence class. The domain of definition of the function can be specified in various ways depending on the nature of the shape variation to be studied. It can be for example, a finite point set, or the real interval, say (0,1).

## 1.3 SOME CRITERIA FOR REPRESENTATION OF SHAPE:

The most useful property of any representation for shape is that it can make some types of information explicit focussing our attention on the essential information, allowing us to look at the shape from another perspective which gives more insight. A shape representation, as discussed in the previous section, must be the same for all the outlines which belong to the same equivalence class.

Another requirement that might be demanded of a shape representation is that it should contain information about the shape at varying levels of detail. Moreover, it should be clear from the representation which features of the shape belong to the coarser levels of detail and which to the finer levels. In other words, the degree of similarity between two shapes must be reflected in their representation, yet, at the same time even subtle differences should be expressible. These two conflicting requirements can be met only if it is possible to decouple stable information that captures the more general and less varying properties within a class from information that is sensitive to the finer variations within the class.

Finally, the representation should be capable of being computed easily.

We shall, at the end, attempt an evaluation of our scheme for representing shape - the so-called Fourier Descriptors - based on the present work and on [ 4 ].

## 1.4 ALGORITHMS FOR SHAPE ANALYSIS:

There are many possible criteria according to which we can classify the methodologies used to analyze shape. First, we can distinguish between the algorithms which examine only the boundary points and ignore the interior area bounded by the boundary, and those which make use of the information contained in the interior region of a boundary as well, e.g. medial axis transformation (MAT). The boundary tracers can, in turn be classified into those which proceed sequentially along the boundary points; and those which examine all or most pairs of boundary points at each stage. Fourier descriptors and parsing of the boundary according to a simple grammar come under the first category while parsing of the boundary according to a context sensitive grammar belongs to the latter class.

Again, we can distinguish between scalar transform techniques, which transform a picture into a set of scalar features e.g., Fourier coefficient and space domain techniques which transform one picture into another, e.g. MAT. An even more fundamental classification is between

information preserving and information non-preserving techniques depending on whether it is possible to reconstruct a reasonable approximation to the original picture from the shape descriptors or not. The reader is referred to [ 8] for a more detailed discussion on the various shape analysis algorithms mentioned above.

## 1.5 PLAN OF THE THESIS:

Chapter 2 gives a brief review of the different approaches that have been adopted in the literature to describe a shape by Fourier methods. The methodology that has been followed by us is then presented in detail. In Chapter 3 we describe the two methods we imployed for classifying closely related shapes. One is based on the concept of bending energy and the other on that of matched filtering. Chapter 4 will deal with the experiments that have been done and the results obtained.

# CHAPTER 2

# SHAPE ANALYSIS AND FOURIER REPRESENTATION

## 2.1 INTRODUCTION:

In Chapter 1, we mentioned some of the schemes for analysing shapes. We shall present here the scheme that we have used for describing shape - the so-called Fourier descriptors. The method of Fourier descriptors constitutes an analytical description of shape in which the principal idea is to convert the description of a geometrical shape to an alternative, equivalent description by means of a linear operator. The 'spatial' domain description is replaced by a 'frequency' domain description.

## 2.2 FOURIER DESCRIPTORS: A REVIEW:

Consider a plane closed curve C in the x,y plane. Such a curve can be Fourier analysed in different ways. The Fourier descriptors of the curve are then defined in terms of the coefficients of the Fourier series expansion. It should be noted that this is an analytical method of describing a curve, an information preserving transformation of the original curve into a set of scaler features. Moreover, one can get as exact a representation as one wishes by simply considering sufficient number of coefficients.

The curve C can be represented by a set of parametric equations in the parameter t

$$x = f_x(t)$$

$$y = f_y(t) \qquad (2.2.1)$$

The function $f_x$ and $f_y$ are periodic in the parameter t and hence can be expressed in a Fourier Series. If we take the parameter t to be the arc-length s along the curve, measured from some arbitrary reference point on the curve, then we have a special form of the description of a curve parameterized by its own arc-length:

$$x = x(s)$$

$$y = y(s) \qquad (2.2.2)$$

Each of the functions x(s), and y(s) can be expressed in a Fourier series:

$$x(s) = a_o + \sum_{n=1}^{\infty} a_n \cos\frac{2n\pi s}{L} + b_n \sin\frac{2n\pi s}{L}$$

$$y(s) = c_o + \sum_{n=1}^{\infty} c_n \cos\frac{2n\pi s}{L} + d_n \sin\frac{2n\pi s}{L} \qquad (2.2.3)$$

where L is the total length of the curve.

The set $a_n, b_n, c_n, d_n$ $n=0,1, \ldots\infty$ of the coefficients is defined to be the Fourier descriptor of the curve. This is the approach that has been followed in this thesis.

Granlund [3], Richard and Hemami [10], Persoon and Fu [9] among others have followed an essentially similar method as above but conceived of the curve as a complex function $\phi(s)$

$$\phi(s) = x(s) + j\, y(s)$$

The Fourier descriptors, $F_n'$ are the coefficients of the complex Fourier series expansion of $\phi(s)$.

$$\phi(s) = \sum_{n=-\infty}^{\infty} F_n' \; e^{+j \frac{2\pi n}{L} s}$$

The curvature K at each point on the curve forms another important descriptor of the shape of a plane curve. The so-called intrinsic equation of a curve expresses the curvature K as a function of the arc length s.

$$K = K(s)$$

The parametric description of a curve, as in Eq. (2.2.1) and implicit description of the curve as $f(x,y) = 0$, contain information about the position, and orientation of the curve relative to the coordinate axes as well as information about the shape of the curve. This is reflected directly in the frequency domain by the dependence of the Fourier coefficients on the position and orientation of the curve in a coordinate system. The intrinsic equation of the curve describes only

the pure shape of the curve without reference to any coordinate axes. Since, moreover, the intrinsic equation of a closed curve is periodic in s i.e., $K(s+L) = K(s)$, we can expand it in a Fourier series. Unfortunately however, for curves which have corners, the curvature is infinite at the corner points and the method fails. The tangent angle $\phi$ at any point on the curve is related to the curvature K through the equation:

$$K = \frac{d\phi}{ds}$$

If we define the cumulative angular function $\theta(s)$ as the difference between the tangent angle at the point s and that at the starting point s=0, then we have

$$\theta(s) = \phi(s) - \phi(o) = \int_o^s K.ds$$

Defining $\theta^*(t)$ as

$$\theta^*(t) = \theta(\frac{Lt}{2\pi}) + t \quad 0 \leq t \leq 2\pi$$

we get a function which is invariant under translation, rotation and scale change. Zahn and Ruskies [12 ], Bennet and MacDonald [ 2], among others, expand $\theta^*(t)$ in a complex Fourier series.

$$\theta^*(t) = \sum_{n=-\infty}^{\infty} F_n \, e^{+jnt}$$

and the set $F_n$ $n = -\infty, \ldots 1, 2, \ldots \infty$ forms the Fourier descriptor of the curve.

It is interesting to compare the different methods of obtaining the Fourier descriptors. The function $\theta^*(t)$ in the tangent angle method is almost always discontinuous and consequently the Fourier coefficients can decrease no faster than $c/n$, where $c$ is a constant independent of $n$, as $n \to \infty$. On the other hand, the parametric functions $x(s)$ and $y(s)$ are always (piece-wise) continuous functions of $s$; hence the coefficients in their Fourier expansion decrease at least as rapidly as $c/n^2$ as $n \to \infty$. Another important difference is that a subset of the descriptors $F_n$ usually does not give a closed curve, whereas a subset of $F_n'$ does. One more advantage of using the coordinates $x(s)$, $y(s)$ instead of the tangent-angle representation of the boundary curve is that the parametric representation is less sensitive to the noise inherent in a fuzzy boundary. The tangent angle is related to the derivatives of $x(s)$ and $y(s)$. A direct consequence of this is that small variations in the coordinate values of the boundary points can result in large variations in the tangent vector.

The Fast Fourier transform or direct integration has been used in the past to calculate Fourier descriptors. However, if the contour is represented by a piece-wise linear approximation, a more direct method of calculating the Fourier coefficients can be employed, as has been shown in Kuhl and Giardina [ 5 ] who also have described a normalization procedure, which we will borrow, based on the elliptic properties of the Fourier coefficients. As mentioned before, the Fourier coefficients are dependent on the choice of the starting point of the traversal round the contour, and on translation, rotation and uniform scaling of the contour.

In many applications in pattern recognition, the position size and rotation of an object are not important. If we describe the object by its boundary, the starting point is irrelevant also. Hence, self-consistent normalization procedures based only on the intrinsic properties of the curve must be used.

Granlund [ 3] considered products of Fourier series coefficients which are invariant to position size, orientation of the contour and starting point choices. The result is an increase in data dimensionality from n to $n^2/2$ without any change in total information. Persoon and Fu [9 ] deal

with unnormalized coefficients and reduce the problem of shape matching to one of finding an optimum size, orientation and starting point match between the sample FD (Fourier descriptors) and each reference FD in the training set. Although this method retains all the shape information of the original contour, and an optimum match for size, orientation, and starting point can always be found, it is computationally expensive. Richard and Hemami [10] follow a similar matching criterion in their work on identification of airplanes.

The invariant elliptic properties of the Fourier coefficients can be exploited for a convenient and intuitively pleasing scheme of normalization, so as to get pure shape information, discarding irrelevant information introduced by the rotation, translation, size and starting point factors.

## 2.3 THE METHOD:

### 2.3.1 Fourier Descriptors of a Piece-wise Linear Contour:

Consider a plane closed vurve C in the x-y plane which is described in terms of the parametric equations

$$x = x(s)$$

$$y = y(s)$$

where s is its own arc-length measured from some arbitrary reference point on the curve. The parameter s varies from

O to L where L is the total length of the curve. The x and y projections are periodic functions in the parameter s.

$$x\,(s+L) = x(s)$$

$$y\,(s+L) = y(s)$$

We can express x(s) and y(s) in a Fourier series:

$$x(s) = A_o + \sum_{n=1}^{\infty} a_n \cos\frac{2n\pi s}{L} + \sum_{n=1}^{\infty} b_n \sin\frac{2n\pi s}{L}$$

$$y(s) = C_o + \sum_{n=1}^{\infty} c_n \cos\frac{2n\pi s}{L} + \sum_{n=1}^{\infty} d_n \sin\frac{2n\pi s}{L} \qquad (2.3.1.1)$$

The coefficients $a_n, b_n, c_n$ and $d_n$ are given by

$$A_o = \frac{1}{L}\int_o^L x(s)ds, \quad c_o = \frac{1}{L}\int_o^L y(s)ds$$

$$a_n = \frac{2}{L}\int_o^L x(s)\cos\frac{2n\pi s}{L}\,ds, \quad c_n = \frac{2}{L}\int_o^L y(s)\cos\frac{2n\pi s}{L}\,ds$$

$$b_n = \frac{2}{L}\int_o^L x(s)\sin\frac{2n\pi s}{L}\,ds, \quad d_n = \frac{2}{L}\int_o^L y(s)\sin\frac{2n\pi s}{L}\,ds \qquad (2.3.1.2)$$

Now suppose that we use a piece-wise linear representation for x(s) and y(s) i.e. we consider C to be described by a set of discrete points $(x_i, y_i)$ i= 0,1,...m . Then, the

derivatives $\frac{dx}{ds}$ and $\frac{dy}{ds}$ consist of a sequence of piece-wise constant derivatives $\Delta x_p / \Delta s_p$ and $\Delta y_p / \Delta s_p$ in the interval $s_{p-1} < s < s_p$ as p varies from 0 to m. We note that the derivative $\frac{dx}{ds}$ is itself periodic and hence can be expanded in a Fourier series :

$$\frac{dx}{ds} = \sum_{n=1}^{\infty} \alpha_n \cos \frac{2n\pi s}{L} + \beta_n \sin \frac{2n\pi s}{L}$$

where

$$\alpha_n = \frac{2}{L} \int_0^L \frac{dx}{ds} \cos \frac{2n\pi s}{L} \, ds$$

$$\beta_n = \frac{2}{L} \int_0^L \frac{dx}{ds} \sin \frac{2n\pi s}{L} \, ds$$

Thus,

$$\alpha_n = \frac{2}{L} \int_0^L \frac{dx}{ds} \cos \frac{2n\pi s}{L} \, ds$$

$$= \frac{2}{L} \sum_{p=0}^{m} \frac{\Delta x_p}{\Delta s_p} \int_{s_{p-1}}^{s_p} \cos \frac{2n\pi s}{L} \, ds$$

$$= \frac{1}{n\pi} \sum_{p=0}^{m} \frac{\Delta x_p}{\Delta s_p} \left( \sin \frac{2n\pi s_p}{L} - \sin \frac{2n\pi s_{p-1}}{L} \right)$$

$$\beta_n = \frac{2}{L} \int_0^L \frac{dx}{ds} \sin \frac{2n\pi s}{L} \, ds$$

$$= \frac{2}{L} \sum_{p=0}^{m} \frac{\Delta x_p}{\Delta s_p} \int_{s_{p-1}}^{s_p} \sin \frac{2n\pi s}{L} \, ds$$

$$= - \frac{1}{n\pi} \sum_{p=o}^{m} \frac{\Delta x_p}{\Delta s_p} \left( \cos \frac{2n\pi s_p}{L} - \cos \frac{2n\pi s_{p-1}}{L} \right)$$

Now, we can obtain $\frac{dx}{ds}$ by directly differentiating the Fourier expansion of x term by term.

$$\frac{dx}{ds} = \sum_{n=1}^{\infty} - \frac{2n\pi}{L} a_n \sin \frac{2n\pi s}{L} + \frac{2n\pi}{L} b_n \cos \frac{2n\pi s}{L}$$

Equating coefficients obtained in the two expressions for $\frac{dx}{ds}$ we get:

$$a_n = \frac{L}{2n^2\pi^2} \sum_{p=o}^{m} \frac{\Delta x_p}{\Delta s_p} \left[ \cos \frac{2n\pi s_p}{L} - \cos \frac{2n\pi s_{p-1}}{L} \right]$$

$$b_n = \frac{L}{2n^2\pi^2} \sum_{p=o}^{m} \frac{\Delta x_p}{\Delta s_p} \left[ \sin \frac{2n\pi s_p}{L} - \sin \frac{2n\pi s_{p-1}}{L} \right]$$

Similarly for the y-projection we get

$$c_n = \frac{L}{2n^2\pi^2} \sum_{p=o}^{m} \frac{\Delta y_p}{\Delta s_p} \left[ \cos \frac{2n\pi s_p}{L} - \cos \frac{2n\pi s_{p-1}}{L} \right]$$

$$d_n = \frac{L}{2n^2\pi^2} \sum_{p=o}^{m} \frac{\Delta y_p}{\Delta s_p} \left[ \sin \frac{2n\pi s_p}{L} - \sin \frac{2n\pi s_{p-1}}{L} \right]$$

2.3.2 Elliptic Properties of the Fourier Coefficients:

The Fourier series representation for the x and y projections can be written as:

$$x(s) = A_o + \sum_{n=1}^{N} X_n$$

$$y(s) = C_o + \sum_{n=1}^{N} Y_n \qquad o \leq s \leq L \qquad (2.3.2.1)$$

where,

$$X_n(s) = a_n \cos \frac{2n\pi s}{L} + b_n \sin \frac{2n\pi s}{L}$$

$$Y_n(s) = c_n \cos \frac{2n\pi s}{L} + d_n \sin \frac{2n\pi s}{L} \qquad (2.3.2.2)$$

For a given n, the points $(X_n, Y_n)$ form an elliptic locus as s varies from O to L. After a little algebra with equation (2.3.2.2) we get

$$\frac{(d_n^2+c_n^2)X_n^2+(a_n^2+b_n^2)Y_n^2-2(a_nc_n+b_nd_n)X_nY_n}{(a_nd_n - b_nc_n)^2} = 1$$

which we recognise to be the equation of an ellipse. The fact that for each n the points $(X_n, Y_n)$ describes an ellipse allows us to interpret in an interesting way the Fourier series representation of a contour. We can conceive of the original contour as being traced out by the phasor addition of rotating phasors, one for each harmonic. Each phasor sweeps out an ellipse, and the phasor associated with the nth harmonic rotates n times as fast as the fundamental.

More important from the point of view of normalization is the fact that the elliptic loci $(X_n, Y_n)$ are independent

of the starting-point of the trace round the contour. To see this let

$$s = s' + s_o.$$

Then we have, from equations

$$X_n(s'+s_o) = a_n \cos \frac{2n\pi(s'+s_o)}{L} + b_n \sin \frac{2n\pi(s'+s_o)}{L}$$

$$Y_n(s'+s_o) = c_n \cos \frac{2n\pi(s'+s_o)}{L} + d_n \sin \frac{2n\pi(s'+s_o)}{L} \qquad (2.3.2.3)$$

This can be rewritten as

$$X_n'(s') = a_n' \cos \frac{2n\pi s'}{L} + b_n' \sin \frac{2n\pi s'}{L}$$

$$Y_n'(s') = c_n' \cos \frac{2n\pi s'}{L} + d_n' \sin \frac{2n\pi s'}{L} \qquad (2.3.2.4)$$

where we have put

$$X_n'(s') = X_n(s'+s_o)$$

$$Y_n'(s') = Y_n(s'+s_o) \qquad (2.3.2.5)$$

The coefficients $a_n'$, $b_n'$, $c_n'$ and $d_n'$ are the coefficients of the expansion for the starting point $s'=0$ i.e., a starting point shifted $s_o$ units along the curve from the original starting point $s=0$. The new set of coefficients $a_n', b_n', c_n', d_n'$ can be obtained from the old set $a_n, b_n, c_n, d_n$ through the transformation T.

$$\begin{bmatrix} a_n' & c_n' \\ b_n' & d_n' \end{bmatrix} = \begin{bmatrix} \cos \frac{2n\pi s_o}{L} & \sin \frac{2n\pi}{L} s_o \\ -\sin \frac{2n\pi s_o}{L} & \cos \frac{2n\pi}{L} s_o \end{bmatrix} \begin{bmatrix} a_n & c_n \\ b_n & d_n \end{bmatrix}$$

(2.3.2.6)

Again, eliminating sines and cosines from equations (2.3.2.4) and using equations (2.3.2.5), (2.3.2.6) we get

$$\frac{(d_n^2+c_n^2)X_n'^2 + (a_n^2+b_n^2)Y_n'^2 - 2X_n'Y_n'(a_nc_n+b_nd_n)}{(a_nd_n - b_nc_n)^2} = 1$$

which gives the same locus as before. Hence the same elliptic loci are obtained for different starting points. It is this invariance that we will exploit for the starting-point normalization.

2.3.3 Rotation of the Contour:

Any rotation of the contour, with respect to the coordinate axes is reflected in the Fourier coefficients. To see the effect of this rotation on the coefficients, assume that the X,Y axes are rotated through a counter clockwise angle $\phi$ into a new set of U,V axes. The new coordinates will then be given by the familiar transformation:

$$\begin{bmatrix} U \\ V \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

The $X_n$ and $Y_n$ projections on the X and Y axes may be written in matrix form as

$$\begin{bmatrix} X_n \\ Y_n \end{bmatrix} = \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos \frac{2n\pi s}{L} \\ \sin \frac{2n\pi s}{L} \end{bmatrix}$$

The projections $U_n$, $V_n$ on the U,V axes will then be

$$\begin{matrix} U_n \\ V_n \end{matrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} X_n \\ Y_n \end{bmatrix}$$

$$= \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos \frac{2n\pi s}{L} \\ \sin \frac{2n\pi s}{L} \end{bmatrix}$$

$$= \begin{bmatrix} a_n'' & b_n'' \\ c_n'' & d_n'' \end{bmatrix} \begin{bmatrix} \cos \frac{2n\pi s}{L} \\ \sin \frac{2n\pi s}{L} \end{bmatrix}$$

where the axially rotated set of coefficients $a_n''$, $b_n''$, $c_n''$, $d_n''$ are given by

$$\begin{bmatrix} a_n'' & b_n'' \\ c_n'' & d_n'' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix}$$

Finally, if we shift the starting point by $s_o$ units and also rotate the X,Y axes by $\phi$ degrees in the counter-clockwise direction, then the combined effect on the Fourier coefficients can be expressed concisely as

$$\begin{bmatrix} a_n'' & b_n'' \\ c_n'' & b_n'' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_n & b_n \\ c_n & d_n \end{bmatrix} \begin{bmatrix} \cos\frac{2n\pi s_o}{L} & -\sin\frac{2n\pi s_o}{L} \\ \sin\frac{2n\pi s_o}{L} & \cos\frac{2n\pi s_o}{L} \end{bmatrix}$$

2.3.4 Uniform Scaling:

Consider the expression for the coefficient $a_n$

$$a_n = \frac{2}{L} \int_o^L x(s) \cos\frac{2n\pi s}{L}\, ds$$

If we put $\frac{2\pi s}{L} = \theta$ we get,

$$a_n = \frac{1}{\pi} \int_o^{2\pi} x(\theta)\cos n\theta\, d\theta$$

If we blow up or [illegible] the original contour by a factor K, we see immediately from the last equation that the coefficient is multiplied by the factor K. Hence, scaling the contour by a factor K scales all the coefficients by the same factor K.

2.3.5 Normalization of the Fourier Descriptors:

We have seen in the last section that the Fourier coefficients contain information about the starting-point, spatial rotation, size and translation of the contour relative to a coordinate system, as well as information about the intrinsic shape of the curve. Hence normalization must be performed to extract the pure shape information. The invariant elliptic properties of the Fourier coefficients form the basis for a convenient method of normalization.

The coefficients $A_o$ and $C_o$ contain the d.c. information or the translation of the contour in the reference frame. By simply ignoring these terms we can achieve independence with respect to translation. The other factors pose a slightly tricky problem. The normalization is done in two stages. To start with, the first harmonic phasor is rotated until it is aligned with a semi-major axis of its locus. This normalizes the starting point. Then the X,Y coordinate axes are swung into a new set of U,V axes defined by the major and minor axes of the ellipse such that the positive X axis is coincident with the semi-major axis located above. This achieves normalization with respect to rotation of the contour.

Consider the first harmonic locus $(X_1, Y_1)$

$$X_1 = a_1 \cos\theta + b_1 \sin\theta$$

$$Y_1 = c_1 \cos\theta + d_1 \sin\theta$$

where $\theta = \frac{2\pi s}{L}$

The starting-point shift $s_o$, or angular shift $\frac{2n\pi s_o}{L}$ can be calculated by finding the value of $\theta$ for which the magnitude of the first harmonic phasor $E = \sqrt{X_1^2 + Y_1^2}$ is a maximum. Differentiating E with respect to $\theta$ and equating it to zero, gives us

$$\theta_o = \frac{1}{2} \tan^{-1} \left[ \frac{2(a_1 b_1 + c_1 d_1)}{a_1^2 + c_1^2 - b_1^2 - d_1^2} \right]$$

That this gives a maximum is checked by evaluating $\frac{d^2E}{d\theta^2}$ at $\theta = \theta_o$ and noting that a negative value is obtained.

The Fourier coefficients, $a_n'$, $b_n'$, $c_n'$, $d_n'$ for the displaced starting point are given by

$$\begin{bmatrix} a_n' & c_n' \\ b_n' & d_n' \end{bmatrix} = \begin{bmatrix} \cos\theta_o & \sin\theta_o \\ -\sin\theta_o & \cos\theta_o \end{bmatrix} \begin{bmatrix} a_n & c_n \\ b_n & d_n \end{bmatrix}$$

To determine the spatial rotation $\phi$ of the contour we write the first-harmonic locus of the point $(X_1'\ Y_1')$

$$X_1'(s') = a_1' \cos \frac{2\pi s'}{L} + b_1' \sin \frac{2\pi s'}{L}$$

$$Y_1'(s') = c_1' \cos \frac{2\pi s'}{L} + d_1' \sin \frac{2\pi s'}{L}$$

Since when s'=0, the first-harmonic phasor is aligned with the semi-major axis, the angle $\phi$ is given by

$$\phi = \tan^{-1} \left[ \frac{Y_1'(o)}{X_1'(o)} \right]$$

$$= \tan^{-1} \left[ \frac{c_1'}{a_1'} \right]$$

The e axially rotated set of coefficients will then be given by:

$$\begin{bmatrix} a_n'' & b_n'' \\ c_n'' & d_n'' \end{bmatrix} = \begin{bmatrix} \cos\phi & \sin\phi \\ -\sin\phi & \cos\phi \end{bmatrix} \begin{bmatrix} a_n' & b_n' \\ c_n' & d_n' \end{bmatrix}$$

If we divide each of the coefficients by the length of the semi-major axis $\sqrt{X_1'(o)^2+Y_1'(o)^2} = \sqrt{(a_1')^2+(c_1')^2}$ then the coefficients are made independent of size as well.

2.3.6 Some Remarks:

The Fourier series representation for shape requires a closed contour. Open outlines have been effectively made closed by retracing the outlines from the end back to the beginning. We are currently working with 12 harmonics as this gives a fairly close (2 pixels) approximation to the original.

# CHAPTER 3

# CLASSIFICATION OF SHAPES USING BENDING-ENERGY AND MATCHED-FILTERING

## 3.1 INTRODUCTION:

An intuitive way of classifying shapes is on the basis of how 'wiggly' they are. Translating this intuition into mathematical terms, we can employ curvature or the rate of change of the tangent with respect to the arc length, as a descriptor of shape. Curvature by itself and measures based upon curvature have played an important role in shape representation and pattern recognition. Indeed, it will not be an exaggeration to say that curvature, either directly or indirectly has played a prominent part in most shape analysis schemes. If we are to believe some of the theories of vision, curvature plays an important role in human perception of shape [ ]. Unfortunately, curvature is too sensitive to noise inherent in a fuzzy boundary to be of much use as a shape descriptor when used by itself. On the other hand, if we average out the curvature over a length of the boundary, the average might serve as an useful measure of how bent or curved the boundary is Interestingly enough, the stored free energy in an elastic bar under deformation is proportional

to $\int K^2 dt$ where K is the curvature and t the length along the bar. Such measures have been used to classify biological shapes [11]. A shape, seen from this point of view is a deformed rod. The bending energy of such a shape can be related to the Fourier coefficients in a simple manner as will be shown in the succeeding section.

Another idea for classifying shapes, based upon energy, that we have pursued is that of optimum filtering or the so-called matched-filtering. We can consider eachs shape as a signal and the problem of shape classifications becomes one of determining how closely the signals clutter together in the signal space.

## 3.2 CLASSIFICATION OF SHAPE USING BENDING ENERGY:

The shape that an elastic bar assumes, when its ends are constrained, is determined by the condition that the free energy associated with the deformation shall be a minimum. If it is assumed that the cross-section is small in comparison with the length of the bar, and further that there is no appreciable torsion of the bar, then the bending energy is proportional to $\int K^2 dt$. Consider a rod of negligible cross-section deformed as shown in Fig. 3.2.1. Then the free energy stored in the rod at a point t on it is given by [6 ]

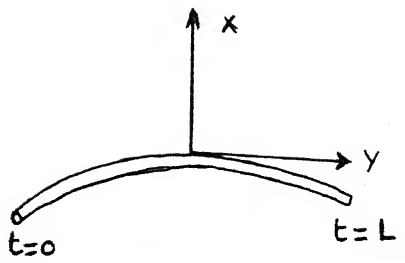



Fig. 3.2.1: An elastic bar under deformation

$$e = \frac{1}{2} Y(t)\ I_y(t)\ K^2(t)$$

where,

$Y(t)$ is the Young's modulus at the point t,

$I_y(t)$ is the moment of inertia of cross-sectional area about the y axis, and

$K(t)$ is the curvature at point t.

The total energy over the rod of length T is

$$E = \frac{1}{2} \int Y(t)\ I_y(t)\ K^2 dt$$

If the cross-sectional dimensions are negligible compared to the length of the rod, then $I_y(t)$ is a constant, say I; if we assume further that the Young's modulus is also a constant Y, then we have

$$E = \frac{1}{2} YI \int_0^T K^2(t)dt$$

If we put $E_T = E/(T/_2 YI)$ we get

$$E_T = \int_0^T K^2 dt \qquad (3.2.1)$$

Now consider a plane closed curve (Fig. 3.2.2) parameterized by its own arc length t. We shall adopt the vector notation as it is more convenient. Vectors will be represented by a symbol with an arrow above, and unit vector will be identified by a cap above the symbol. Referring to Fig. 3.2.2, the position vector $\vec{R}$ is given by

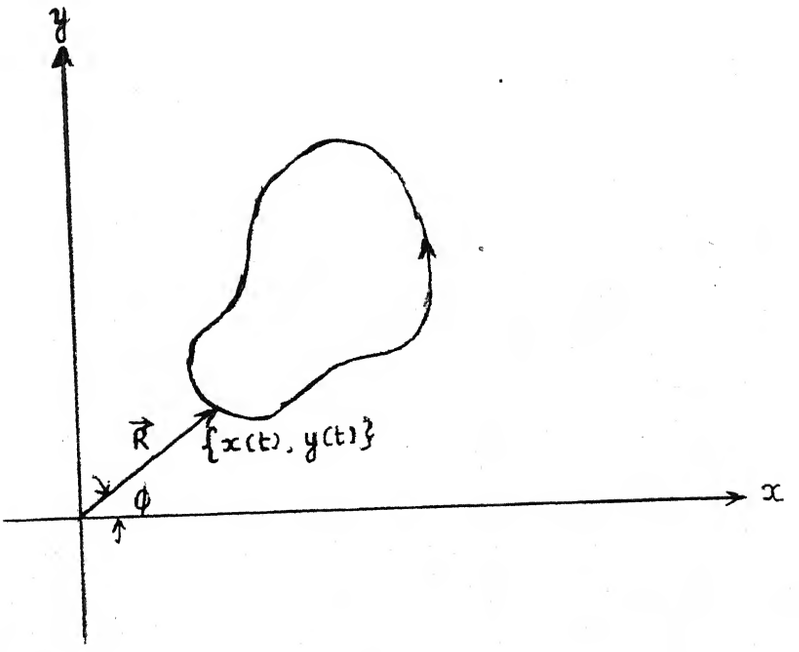



Fig. 3.2.2: Calculation of bending energy of a shape.

$$\vec{R} = x(t)\,\hat{a}_x + y(t)\,\hat{a}_y$$

where $\hat{a}_x$ and $\hat{a}_y$ are unit vectors along the x and y axis respectively. The unit tangent vector T will be then given by

$$\hat{T} = \frac{\vec{dR}}{dt} = \frac{dx}{dt}\,\hat{a}_x + \frac{dy}{dt}\,\hat{a}_y$$

The rate of change of the tangent vector with respect to the arc length is

$$\frac{\vec{dT}}{dt} = \frac{d^2x}{dt^2}\,\hat{a}_x + \frac{d^2y}{dt^2}\,\hat{a}_y$$

But we know that,

$$\frac{\vec{dT}}{dt} = \frac{\vec{dT}}{d\phi} \cdot \frac{d\phi}{dt} = \hat{n}\,K(t)$$

where,

$\hat{n} = \frac{\vec{dT}}{d\phi}$ is the unit normal to the curve

$K(t) = \frac{d\phi}{dt}$ is the curvature at point t

Hence,

$$\left|\frac{dT}{dt}\right|^2 = K^2(t)$$

$$= \left(\frac{d^2x}{dt^2}\right)^2 + \left(\frac{d^2y}{dt^2}\right)^2 \qquad (3.2.2)$$

Substituting in equation (3.2.1) we get,

$$E_T = \int_0^T [(\frac{d^2x}{dt^2})^2 + (\frac{d^2y}{dt^2})^2 ]dt \qquad (3.2.3)$$

Now, consider the Fourier series expansion for x(t) and y(t).

$$x(t) = A_o + \sum_{n=1}^{\infty} a_n \cos n\omega_o t + b_n \sin n\omega_o t$$

$$y(t) = C_o + \sum_{n=1}^{\infty} c_n \cos n\omega_o t + d_n \sin n\omega_o t \qquad (3.2.4)$$

where $\omega_o = \frac{2\pi}{T}$ .

Parseval's relation states that for any signal f(t) represented by a Fourier series

$$f(t) = A_o + \sum_{n=1}^{\infty} a_n \cos n\omega_o t + b_n \sin n\omega_o t$$

the following relation is true:

$$\frac{1}{T} \int_0^T f^2(t)dt = A_o^2 + \frac{1}{2} \sum_{n=1}^{\infty} (a_n^2 + b_n^2).$$

From eqns. (3.2.3) and (3.2.4) and using Parseval's relation we get,

$$E_T = T \cdot \frac{1}{2} \sum_{n=1}^{\infty} (n\omega_o)^4 (a_n^2+b_n^2+c_n^2+d_n^2) \qquad (3.2.5)$$

If we define an average bending energy per unit length then we get

$$E_N = \frac{E_T}{T} \cdot$$

$$= \frac{1}{2} \sum_{n=1}^{\infty} (n\omega_o)^4 \, (a_n^2 + b_n^2 + c_n^2 + d_n^2) \qquad (3.2.6)$$

A little clarification with regard to the last equation is in order. If we are to leave the coefficients $a_n, b_n, c_n$ and $d_n$ unnormalized, then the bending energy $E_N$ will not be the same for an outline and an uniformly scaled version of it. If, for example, we blow up the outline by a factor m, the magnified outline will have a bending energy that is lesser by a factor $m^2$. Although this observation is what we should intuitively e expect, yet we need a shape factor that is identical for all boundaries in the same equivalence class. On the other hand, if we choose to use normalized coefficients, we get a shape factor that is the same for all related, translated, and scaled versions of the original outline. The normalized coefficients, thus, fit in very neatly in this scheme.

Equation (3.2.6) reveals some very interesting features. The bending energy is proportional to a weighted sum of the squares of the Fourier coefficients. The presence of the factor $n^4$ implies that as n increases, the weight given to the

harmonic coefficients increases very rapidly. The higher harmonics are given far more importance than the lower harmonics. For example, the relative importance ascribed to the first harmonic and the tenth harmonic will be in the ratio $1:10^4$. Now, we intuitively expect that the lower harmonics determine the gross features of the shape while the higher harmonics control the finer details. Pattern recognition work done by Karnick [ 4], Fu [ 9] and others provide experimental evidence to support such an hypothesis. Keeping this fact in mind, then, our scheme of weighted sum with the weights increasing exponentially with n seems well suited for distinguishing between shapes whose general pattern is unvarying but which exhibit subtle differences between each other. It will be seen in the next chapter that such indeed is the case. In such a classification, all shapes with identical stored energy in their contours will belong to the same equivalence class.

It is possible to make still finer classifications of shapes by considering the energy stored over small lengths of the boundary, instead of the total energy over the whole permiter. Indeed, the incremental nergy stored in the boundary, between $t_1$ and $t_2$ will be

$$\Delta E = \int_{t_1}^{t_2} K^2(t)dt \qquad (3.2.7)$$

From equations (3.2.2), (3.2.4) and (3.2.7), we get,

$$\Delta E = \int_{t_1}^{t_2} [(\frac{d^2x}{dt^2})^2 + (\frac{d^2y}{dt^2})^2]dt$$

$$= \int_{t_1}^{t_2} ( \sum_{n=1}^{N} (n\omega_o)^2 [a_n \cos n\omega_o t + b_n \sin n\omega_o t])^2 dt$$

$$+ ( \sum_{n=1}^{N} (n\omega_o)^2 [c_n \cos n\omega_o t + d_n \sin n\omega_o t])^2 dt$$

(3.2.8)

The RHS of equation (3.2.8) can be integrated using numerical methods.

To illustrate the usefulness of such an inv incremental measure suppose that we are required to distinguish between a shape and a reflected image of the shape (see Fig. 3.2.3). The total energy in the shape will be the same for the shape and its mirror image. However if both the contours are traversed in the same direction, assuming that we start the traversals at the points marked, then, it is obvious that the energies stored in the lengths ab and a'b' will be different.

## 3.3 CLASSIFICATION OF SHAPE USING MATCHED FILTER:

We can consider the issue of recognizing a shape as that of detecting a signal in the presence of additive noise.

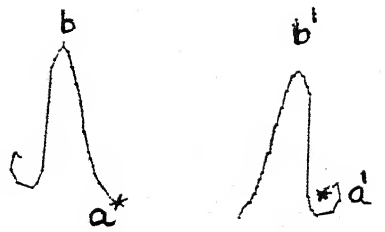



Fig. 3.2.3: A shape and its mirror image

In detection, one is concerned not so much with reproducing the original signal as with being able to establish the presence or absence of the signal. Typically, the form of the signal is known in advance and one is required to decide whether the signal component is present in the received wave, corrupted by additive noise. It turns out that when the power spectral density of the noise is constant i.e. when the noise is white, the optimum solution to the problem of detection is a matched filter. The detection system in such a case is called a matched filter because its characterization is matched to that of the signal component, in the received complex of signal plus noise. A matched filter is optimum in the sense that it maximizes the signal to noise ratio at the filter output; this ratio depends only upon the ratio of the signal energy to the spectral density of the white noise at the filter input.

The Fourier series representation for $x(t)$ and $y(t)$ suggest that the signals $x(t)$ and $y(t)$ can be considered to be vectors in the infinite dimensional vector space in which the set $\cos n\omega_o t$, $\sin n\omega_o t$; $n = 1,2,\ldots$ form an orthogonal basis set. We know, however, that in practice all signals must have a finite energy and hence thecoefficients $a_n$ and $b_n$ must tend to zero for large values of n. Hence, although the

dimensionality of the signal space is infinite we can, for all practical purposes consider an approximate finite dimensional space with dimension N. Any signal, then, can be considered to be a point in this N-dimensional space. Viewed in this perspective, optimum filtering turns out essentially to be a scheme wherein, given a set of m signal points $s_i(t)$ $i = 1,...m$ in the signal space, the unknown signal is declared to be identical to that signal for which the distance from the unknown signal is a minimum [ 7 ].

The received signal $s_r$ is declared to be that signal $s_i$, for which the decision function

$$D = \left| \sum_{n=1}^{\infty} (a_{in}^2 + b_{in}^2) - \sum_{n=1}^{\infty} (a_{in}\, a_{rn} + b_{in}\, b_{rn}) \right| \qquad (3.3.1)$$

is a minimum. It should be noted that when i=r, D=o.

# CHAPTER 4

# EXPERIMENTS, RESULTS AND DISCUSSION

## 4.1 INTRODUCTION:

As we stated in the beginning, one of our main aims was to investigate the degree to which the Fourier series representation for shape can achieve decoupling, in the representation, between the gross and fine features of a shape and we stated in the previous chapter our belief that the lower harmonics dictate the gross features of the shape and the higher harmonics the finer features. In Chapter 3, we also argued out how, based on the above assumption, the bending energy associated with a shape constitutes a good basis for sorting out variations within a class. We also pointed out how classifying shapes within a class may be considered within the framework of matched filtering of signals.

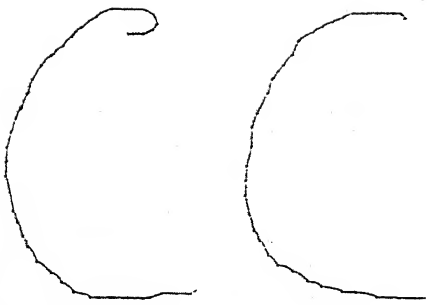
## 4.2 THE GLOBAL VERSUS LOCAL NATURE OF THE FOURIER REPRESENTATION:

As seen from a purely mathematical point of view, the Fourier coefficients have the property that of all trigonometric polynomials of order n, that polynomial which has as its coefficients the Fourier coefficients of the function f(x), has the least mean-square deviation from that function f(x).

The Fourier series of a function f(x) converges to the function f(x) at all points where f(x) is continuous and to the mean value of the jump where there is a jump discontinuity. Since the coefficients of the Fourier series are obtained by integration, any change in any part of the curve will affect all the coefficients. While, it is true that a drastic change such as a step discontinuity introduced over a small portion of the curve, will adversely affect the convergence of the series everywhere (the series will now converge more slowly), any small change in the function values in a narrow neighbourhood which does not destroy the differentiability of the function will affect the higher harmonics more than the lower harmonics.

To what extent our assertion that the lower harmonics control the course of the function in a global sense while the higher harmonics are responsible for local variations can be seen from figures 4.2.1 and 4.2.2. Strokes, lobes, curls etc. are regions of high curvature and contribute to the presence of significant higher harmonics in the Fourier representation.

Part (a) of Fig. 4.2.1 shows a shape generated with all the twelve coefficients being taken into account. Part (b) shows the figure obtained by considering only the first four harmonics.



(a) (b)

Fig. 4.2.1: Illustrating the role of lower and higher harmonics.

a) Shape re-generated with all the 12 harmonics
b) Shape regenerated by considering only the first four harmonics.

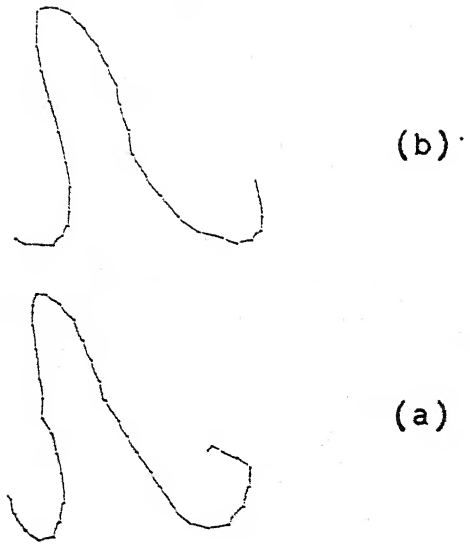



Fig. 4.2.2: Illustrating the role of lower and higher harmonics.

a) Shape regenerated with all the 12 harmonics

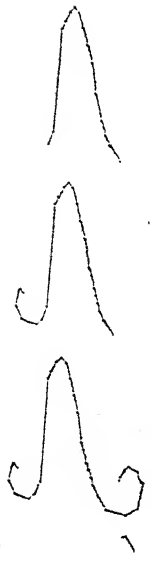
b) 8-harmonic approximation

It is evident from the above examples that our contention that there is a decoupling, fuzzy though it be, between the gross and fine features of a shape, in terms of the lower and higher harmonics, is quite admissible. Moreover, Karnick [4 ] has found that reliable interclass separation of hand-written characters can be achieved by considering only the first five harmonics. This corroborates our claim that the more stable and less varying properties of a shape are captured by a few lower harmonics.

## 4.3 CLASSIFICATION BASED ON BENDING-ENERGY:

In Section 3.2, we saw how a boundary or outline of an object can be considered to be an elastic bar undergoing deformation. We also saw how the free energy stored in the shape can be related to the Fourier coefficients via equation (3.2.6). This point of view leads to the use of the bending energy associated with a shape to generate a set of equivalence classes, all shapes within a class being equivalent in terms of the stored energy in them.

Fig. 4.3.1 shows three shapes which are closely allied to each other and the bending energy associated with them. It should be noted that the classification is in harmony with the way an human observer would classify them, or based on his intuitive ideas of how curved 'or' 'wiggly' a shape is.



B.E. = 288.805

B.E. = 530.897

B.E. = 766.118

Fig. 4.3.1: Three shapes and their bending energies.

Some more shapes and their bending energies are given in Fig. 4.3.2.

Fig. 4.3.3 shows fifty one variants of the character 'C' arranged in the descending order based upon bending energy. The values of the bending energies are listed in Table 4.3.1.

We see from Fig. 4.3.3 how the smoother variants - those devoid of any strokes, lobes, curls etc. cluster together at the lower end of the energy spectrum and how the more wiggly ones occupy the higher end of the spectrum. It should be pointed out further that shapes which are quite close to each other are also classified at the same level. For example, the variants C<22>, C<26> and C<27> are quite close to each other and this is directly reflected in terms of their bending energies. It is also a vindication of the fact that the Fourier series representation is stable - the degree of similarity between two shapes is reflected in the representation as well, which is sufficiently insensitive to minor variations.

The results of similar classifications performed on the set of 51 variants of the character W and S are given in Tables 4.3.2 and 4.3.3 respectively. The reader is referred to the Appendix for a list of the figures.

B.E. = 83.838546

B.E. = 70.259707

B.E. = 465.14971

Fig. 4.3.2: Bending energies of three variants of B

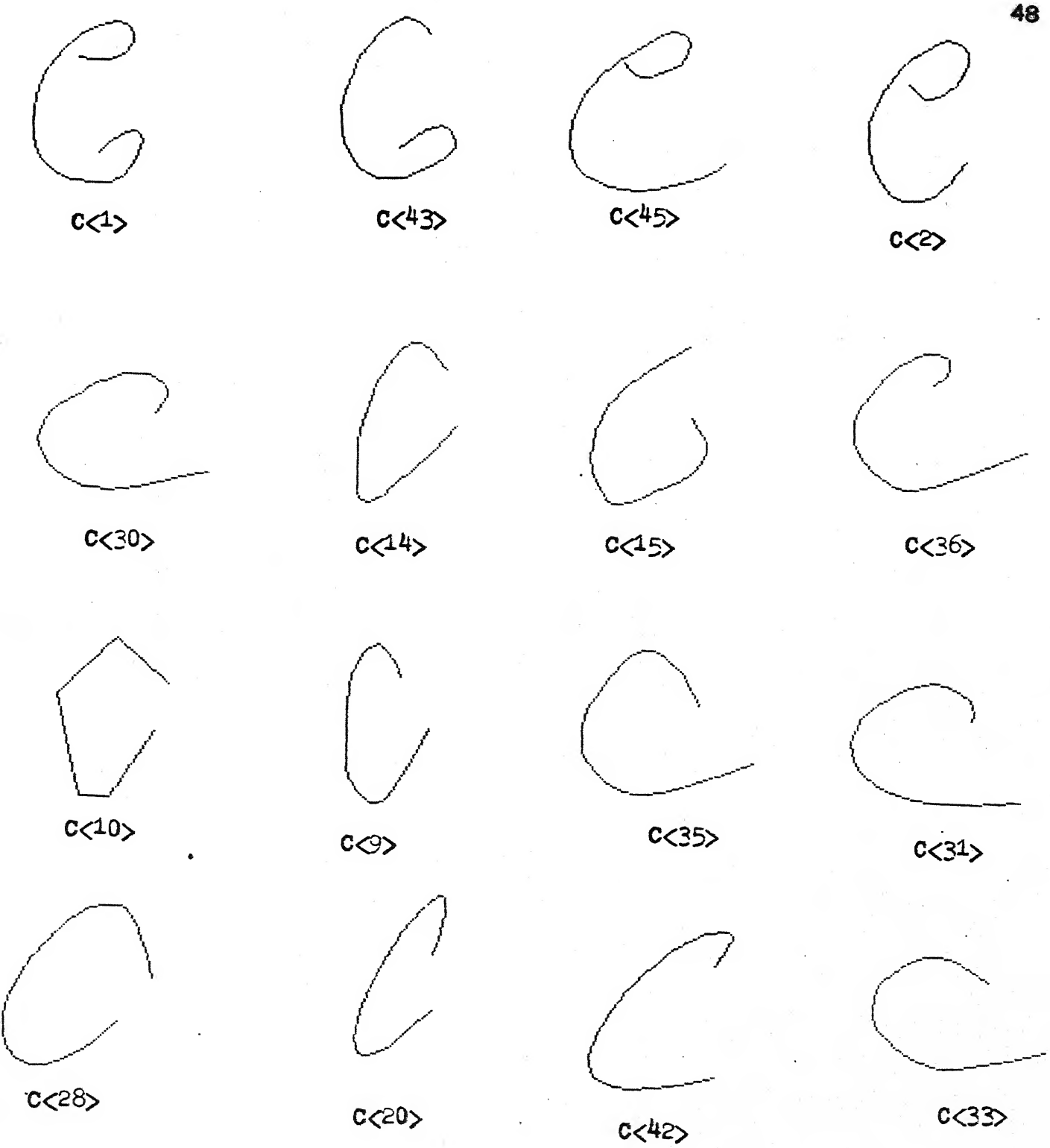
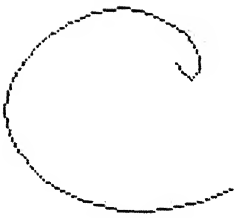



Fig. 4.3.3: 51 Variants of 'C' arranged in Descending Order of Bending Energy.

C<21>



C<34>

C<7>

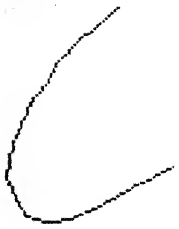
C<32>

C<17>

C<0>

C<37>

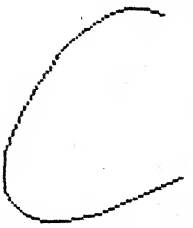
C<49>



C<16>

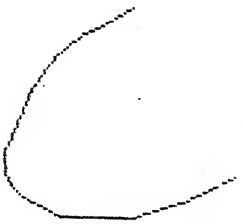
C<41>

C<8>



C<23>

C<24>



C<26>

C<22>

C<27>

Fig. 4.3.3: continued.

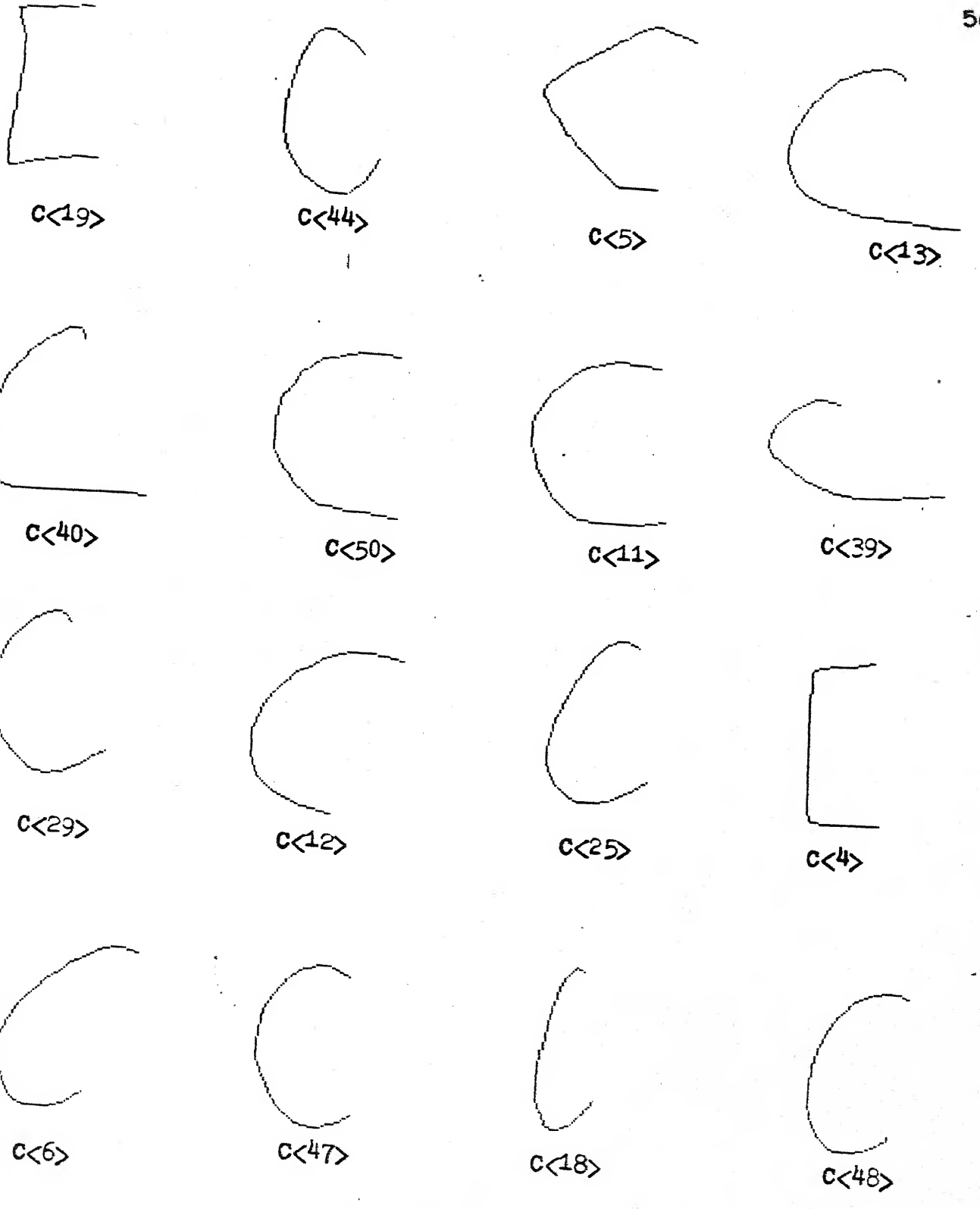



Fig. 4.3.3: continued.

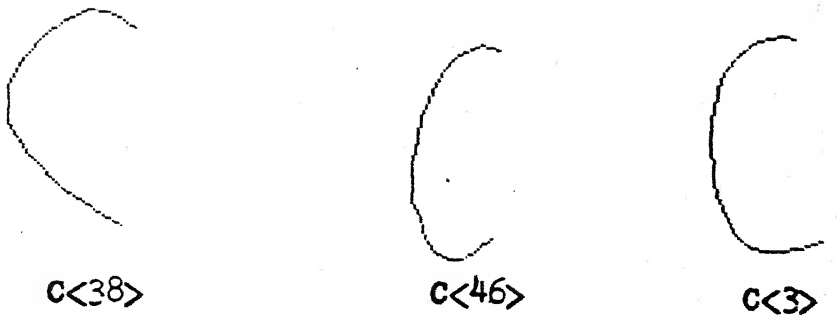



Fig. 4.3.3: continued

Table 4.3.1

The bending energies of 51 variants of 'C' arranged in descending order

| Variant | Bending Energy |
|---|---|
| 1 | 160.75951 |
| 43 | 88.822914 |
| 45 | 86.054553 |
| 2 | 83.840840 |
| 30 | 80.947044 |
| 14 | 69.530754 |
| 15 | 69.527396 |
| 36 | 66.638730 |
| 10 | 61.981963 |
| 9 | 59.903209 |
| 35 | 59.690878 |
| 31 | 54.106864 |
| 28 | 52.050690 |
| 20 | 50.768181 |
| 42 | 49.461648 |
| 33 | 45.966666 |
| 21 | 45.281404 |
| 34 | 43.955914 |
| 7 | 41.664268 |
| 32 | 41.302578 |
| 17 | 40.946887 |
| 0 | 37.394480 |
| 37 | 34.373846 |
| 49 | 32.881071 |
| 16 | 32.221905 |
| 41 | 31.470717 |

| Variant | Bending Energy |
|---|---|
| 8 | 29.725692 |
| 23 | 29.585152 |
| 24 | 29.441790 |
| 26 | 28.719499 |
| 22 | 28.086087 |
| 27 | 27.909746 |
| 19 | 26.613102 |
| 44 | 25.322420 |
| 5 | 25.186769 |
| 13 | 23.717210 |
| 40 | 21.864629 |
| 50 | 21.689877 |
| 11 | 21.093493 |
| 39 | 19.772351 |
| 29 | 19.004930 |
| 12 | 18.531553 |
| 25 | 18.424521 |
| 4 | 17.905297 |
| 6 | 17.671253 |
| 47 | 17.572102 |
| 18 | 17.162462 |
| 48 | 16.792473 |
| 38 | 15.712571 |
| 46 | 14.330127 |
| 3 | 14.270069 |

Table 4.3.2: The bending energies of 51 variants of 'W' arranged in descending order.

| Variant | Bending energy |
|---|---|
| 45 | 805.92713 |
| 0 | 571.74282 |
| 12 | 383.86132 |
| 49 | 382.55382 |
| 15 | 367.47211 |
| 48 | 294.20727 |
| 46 | 286.87660 |
| 3 | 229.65457 |
| 40 | 207.92111 |
| 38 | 205.52481 |
| 36 | 199.04057 |
| 50 | 183.43976 |
| 9 | 181.18240 |
| 11 | 176.17072 |
| 7 | 169.44450 |
| 18 | 168.02066 |
| 21 | 163.79804 |
| 28 | 162.33990 |
| 37 | 160.06561 |
| 27 | 158.76482 |
| 14 | 155.52497 |
| 25 | 152.86746 |
| 23 | 152.46783 |
| 10 | 152.15717 |
| 41 | 149.29733 |
| 34 | 144.04170 |
| 47 | 143.48608 |

| Variant | Bending energy |
|---|---|
| 44 | 124.36388 |
| 13 | 117.90828 |
| 35 | 116.69977 |
| 33 | 107.86908 |
| 29 | 104.77307 |
| 19 | 102.62156 |
| 39 | 98.361093 |
| 30 | 94.112065 |
| 32 | 90.154442 |
| 1 | 89.851562 |
| 42 | 88.846672 |
| 17 | 86.225944 |
| 20 | 82.617205 |
| 26 | 79.992243 |
| 16 | 78.948387 |
| 2 | 78.250048 |
| 5 | 72.891058 |
| 22 | 69.170465 |
| 43 | 66.820249 |
| 6 | 64.118934 |
| 31 | 59.640653 |
| 8 | 58.458732 |
| 4 | 56.992058 |
| 24 | 44.251450 |

Table 4.3.3

The bending energies of 51 variants of 'S' arranged in descending order

| Variant | Bending energy |
|---|---|
| 42 | 660.73912 |
| 13 | 159.05835 |
| 39 | 142.65612 |
| 40 | 140.84575 |
| 16 | 129.67440 |
| 37 | 122.98946 |
| 15 | 122.71067 |
| 38 | 113.83172 |
| 17 | 86.963529 |
| 47 | 85.963388 |
| 46 | 81.467459 |
| 1 | 81.042984 |
| 2 | 80.037961 |
| 49 | 69.380047 |
| 44 | 68.725529 |
| 43 | 60.638382 |
| 41 | 59.256546 |
| 0 | 58.791855 |
| 7 | 54.907763 |
| 32 | 53.826262 |
| 30 | 53.520165 |
| 45 | 51.798248 |
| 14 | 51.957600 |
| 50 | 44.494440 |
| 3 | 44.239019 |
| 29 | 42.548772 |
| 5 | 39.555336 |

| Variant | Bending energy |
|---|---|
| 36 | 39.045345 |
| 31 | 38.806584 |
| 35 | 38.039027 |
| 4 | 36.563881 |
| 10 | 35.033606 |
| 27 | 34.559309 |
| 19 | 34.180757 |
| 6 | 33.972211 |
| 26 | 31.898164 |
| 28 | 30.677968 |
| 9 | 30.003176 |
| 18 | 29.004076 |
| 22 | 28.521631 |
| 11 | 28.394356 |
| 21 | 23.058547 |
| 12 | 22.806727 |
| 25 | 21.697493 |
| 48 | 21.649204 |
| 8 | 21.421442 |
| 33 | 19.400690 |
| 34 | 18.255912 |
| 20 | 15.821547 |
| 24 | 14.628535 |
| 23 | 14.524376 |

It is possible to make finer classifications if we employ the strategy of computing energies stored in smaller lengths of the boundary, as discussed in section 3.2, instead of the total energy over the whole length. To explain the philosophy of the method, we pose the problem of distinguishing between a shape and a mirror image of it. Obviously, a shape and a mirror image will have the same total bending energy.

Fig. 4.3.4 shows a shape and its mirror image. The energies stored in the corresponding parts of the two shapes, calculated using eq. (3.2.7) is given in Table 4.3.4. It is evident from Table 4.3.4 that although the total energy in the two shapes are the same, yet a distinction can still be achieved on the basis of energies associated with corresponding parts.

The drastic reduction in the dimensionality of the data from n to 1 makes this method too insensitive to be suitable for obtaining reliable classifications in the larger space containing all hand-written characters. However, if we restrict our attention to the subspace containing the variants of a particular character, then the bending energy measure provides an intuitively appealing basis for discriminating between closely related shapes.

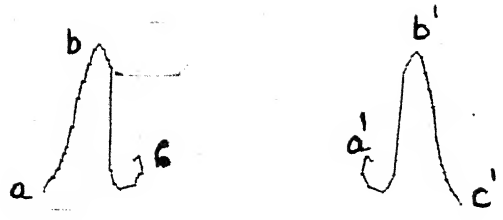



Fig. 4.3.4: A shape and its Mirror Image.

Table 4.3.4: Bending energies of a shape and its mirror image

| Original shape | | Mirror Image | |
|---|---|---|---|
| Length of the curve | Bending Energy | Length of the curve | Bending energy |
| ab | 288.80488 ÷ 2 | a'b' | 383.05923 |
| bc | 383.05923 | b'c' | (288.80488) ÷ 2 |
| Total | 527.46167 | Total: | 527.46167 |

## 4.4 CLASSIFICATION BASED ON MATCHED-FILTERING:

In this approach, we choose a prototype and decide whether a given shape is an instance of the prototype by testing if the decision function D given in equation (3.3.1) is zero or not. In order to classify shapes we assume a certain threshold, fixed at a certain percentage of the energy of the prototype, and put those shapes for which the decision function D is less than the threshold in the same equivalence class.

The first prototype we chose was C<2>. All the 51 variants of the character C were passed through the filter matched to the prototype C<2>. Table 4.4.1 shows the classes obtained when the threshold was fixed at 10% of the energy of the prototype and at 20%.

The next prototype we chose was C<3>. The results obtained in this case are shown in Table 4.4.2.

The prototypes were chosen based on the results of our experiment using the bending-energy concept. A look at table 4.3.1 shows that these two prototypes are widely separated in their energy spectrum - C<3> lies at the bottom of the range while C<2> lies at the top. This observation leads us to expect that the sets of the equivalence classes obtained in the two cases should be disjoint. We find that such indeed, is the case from Tables 4.4.1 and 4.4.2.

Table 4.4.1

RESULTS OF MATCHED-FILTERING

Prototype : C <2>

| % Threshold | Equivalence class obtained |
|---|---|
| 10% | 2, 45, 34, 28 |
| 20% | 2, 45, 34, 28, 10, 17, 0, 42, 20, 41, 5. |

Table 4.4.2
RESULTS OF MATCHED-FILTERING

Prototype : C <3>

| % Threshold | Equivalence class obtained |
|---|---|
| 10% | 3, 19, 18, 39, 47, 44, 45, 1, 29, 24, 49, 9, 50, 11, 43, 40, 13. |

# CHAPTER 5

## CONCLUSIONS AND FURTHER WORK

Two rather conflicting requirements that are demanded of any scheme for representing shape are stability and sensitivity. A representation is stable if insignificant, noise-like variations between two similar shapes do not cause the corresponding representations for them to diverge widely. Indeed, stability is essential for the representation to be used as a basis on which to classify shapes, in as much as it is a measure of correspondance in the representation between two similar shapes. On the other hand, the representation should be capable of expressing subtle but significant or meaningful differences encountered between otherwise similar shapes. In other words, the ease with which one can make classifications within a class of (grossly) similar shapes is a measure of the ability of the representation to capture really fine features of the shape. We examined these issues in the context of the Fourier representation for shape. We have found that such fine distinctions within a class is, indeed, possible by using the intuitively satisfying concept of bending-energy, which is related to the Fourier coefficients in a simple way.

Another useful idea which we explored in this regard is that of considering a shape as a signal and employing matched-filtering to classify shapes, now considered as signals.

In the context of matched filtering, an idea that can be investigated and can prove to be more informative is that of the so-called ambiguity function used in radars. A plot of the ambiguity function - the ambiguity diagram - represents the response of the matched filter to the signal for which it is matched as well as to the mismatched signals. The ambiguity function is a function of two variables: time-delay and frequency. Thus, it is expected that the ambiguity function will prove to be more de descriptive than the approach we have followed.

Another question that crops up in judging a representation is: what class of objects are representable? Indeed any physically realizable shape is capable of being described by a Fourier series representation. One of the objections raised against Fourier descriptors is that for shapes with 'corners' the Fourier series converges rather slowly and hence the number of significant harmonics one should consider to get a specified accuracy of approximation will become large. The problem becomes more accute when one considers a

class of mixed shapes, some of which are smooth (and hence require a lesser number of harmonics) and some of which are not so smooth. In such a case, choosing a larger number of harmonics might lead to an inefficient solution, from the point of view of the time required to calculate the additional coefficients, especially if the shapes which have corners are few. However, there is a way out: the rate at which the coefficients of the Fourier series expansion converge to zero for large n, contains information as to the degree of smoothness of the curve.

If a function f(t) and its first k-1 derivatives satisfy the Dirichlet conditions and are everywhere continuous, then as n becomes infinite the coefficients $a_n$, and $b_n$ in the Fourier series of f(t) tend to zero at least as rapidly as $c/n^{k+1}$ where c is a constant independent of n. If in addition, the kth derivative of f(t) is not everywhere continuous, then either $a_n$ or $b_n$, and in general both, can tend to zero no faster than $c/n^{k+1}$.

Conversely, by observing the rate at which the term in the Fourier series of an otherwise unknown function approach zero, we can obtain useful information about the degree of smoothness of the function, and hence decide whether it would be worth our while to consider more harmonics. It is

## REFERENCES


1. Attneave, F., 'On some informational aspects of vision', Psych. Rev. 61, 183-193, (1954).

2. Bennet, J.R., MacDonald, J.S., 'On the measurement of curvature in a quantized environment', IEEE Trans. C-24, 803-820 (1975).

3. Granlund, G.H., 'Fourier preprocessing for hand-printed character recognition', IEEE Trans. C-21, 195-201, (1972).

4. Karnick, H.C., 'On learning and recognizing patterns with natural variations', Ph.D. Thesis, Indian Institute of Technology, Kanpur, India (1983).

5. Kuhl, F.P., Giardina, F.W., 'Elliptic Fourier Features of a closed contour', Comp. Graphics and Image Proc. 18, 236-258, (1982).

6. Landau, L.D., Lifschitz, E.M., 'Theory of Elasticity' (Perganon, 1959).

7. Lathi, B.P., 'Modern digital and analog communication systems', (Holt-Saunders, 1983).

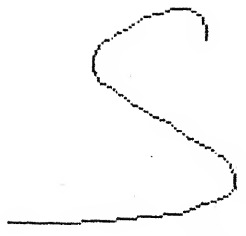
8. Pavlidis, T., 'Algorithms for shape analysis of contours and waveforms', IEEE Trans. PAMI-2, 301-312 (1980).

9. Persoon, E., Fu, K.S., 'Shape discrimination using Fourier Descriptors', IEEE Trans. SMC-7, 170-179 (1977).

10. Richards, C.W., Hemani, H., 'Identification of 3D-objects using Fourier Descriptors of the boundary curve', IEEE Trans. SMC-4, 371-378 (1974).

11. Young, I.T., Walker, J.E., Bowie, J.E., 'A technique for analysis of biological shapes', Inform. Control. 25, 357-370 (1974).

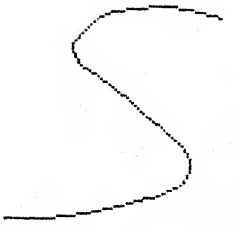
12. Zahn, C.T., Roskies, Z., 'Fourier Descriptors for plane closed curves', IEEE Trans. SMC-7, 170-179, (1977).



S<0>

S<1>

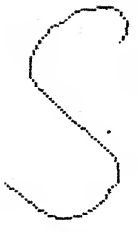
S<2>



S<3>

S<4>

S<5>



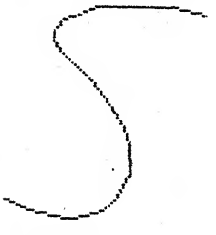
S<6>

S<7>

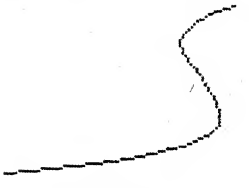
S<8>

S<9>

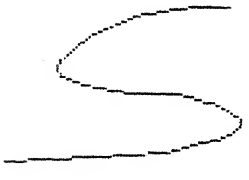
S<10>



S<11>



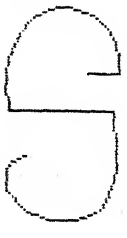
S<12>



S<13>

S<14>

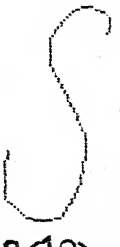
S<15>



S<16>

S<17>

S<18>



S<19>

S<20>

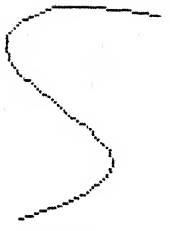
S<21>

S<22>

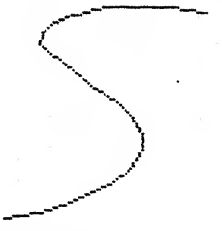
S<23>

S<24>

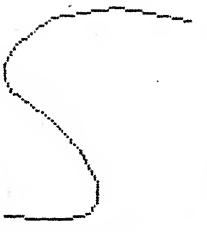
S<25>



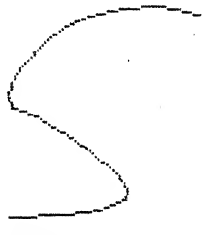
S<26>



S<27>



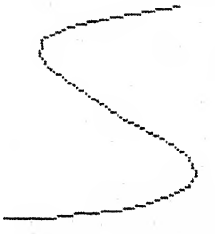
S<28>



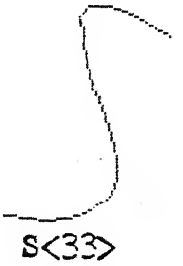
S<29>

S<30>

S<31>

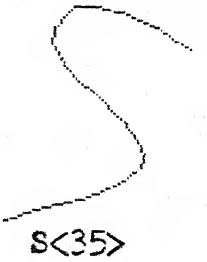


S<32>



S<33>

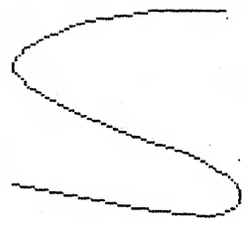
S<34>



S<35>

S<36>

S<37>

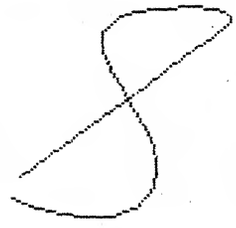


S<38>

S<39>

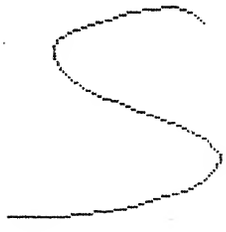
S<40>

S<41>



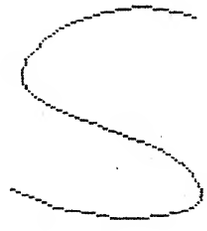
S<42>

S<43>



S<44>

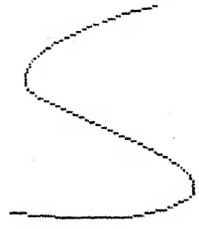
S<45>



S<46>

S<47>

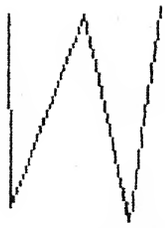
S<48>



S<49>

S<50>

A-4



W<0>

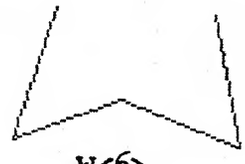
W<1>

W<2>

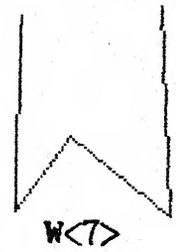
W<3>

W<4>

W<5>

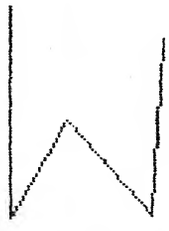


W<6>



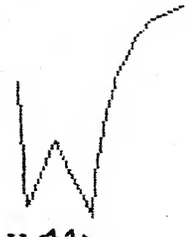
W<7>

W<8>



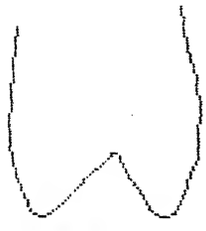
W<9>

W<10>

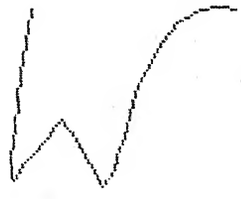


W<11>

W<12>



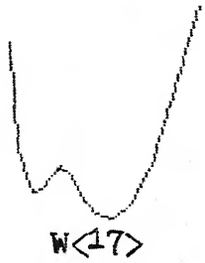
W<13>



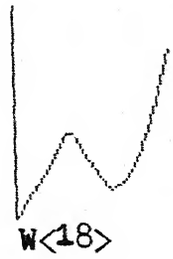
W<14>

W<15>

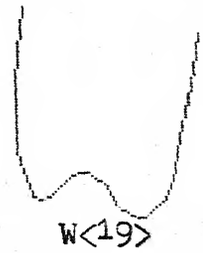
W<16>



W<17>



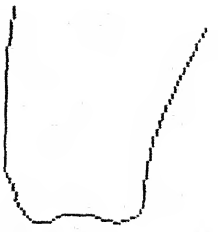
W<18>



W<19>

W<20>

W<21>

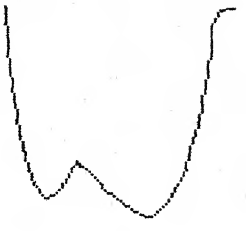


W<22>

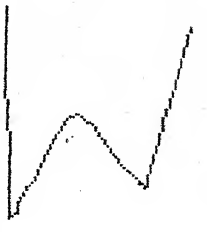
W<23>

W<24>

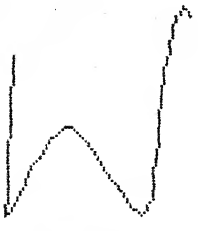
W<25>



W<26>

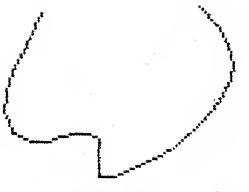


W<27>



W<28>

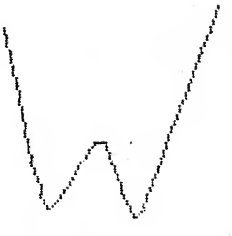
W<30>



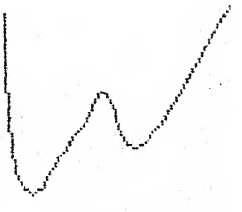
W<31>

W<32>

W<33>



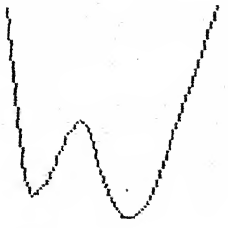
W<34>



W<35>

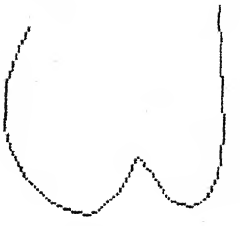
A-6

W<36>

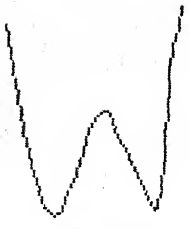


W<37>

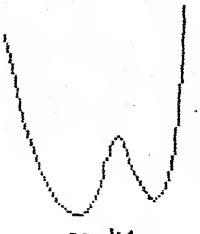
W<38>



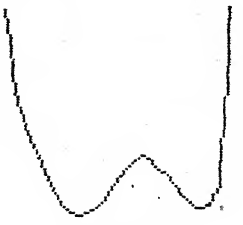
W<39>



W<40>

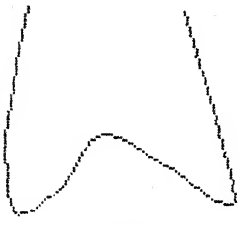


W<41>

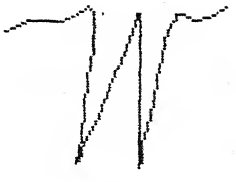


W<42>

W<43>



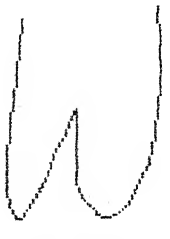
W<44>



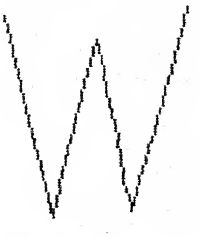
W<45>

W<46>

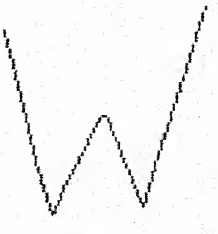
W<47>



W<48>



W<49>



W<50>